Economia Aplicada, v. 21, n. 2, 2017, pp. 285-309

GASTOS EM EDUCAÇÃO: MAIS RECURSOS SEM GESTÃO?

Carlos Renato de Melo Castro *
Geraldo da Silva e Souza †
Maria Eduarda Tannuri-Pianto ‡

Resumo

Neste trabalho, modela-se uma função custo para o ensino fundamental brasileiro. Metodologicamente, destaca-se a utilização da Forma Flexível de Fourier para transpor os vieses normalmente apresentados na utilização das formas flexíveis locais. Os resultados indicam que, tudo mais constante, mesmo gastos da ordem de 10% do PIB não seriam suficientes para alcançar a proficiência mínima para 100% dos alunos. Observou-se que municípios com mais renda, menos urbanos, com menos desigualdade de renda e com maiores taxas de analfabetismo apresentam maiores exigências de gastos e maiores níveis de ineficiência. Quanto às variáveis que estão sob a gestão educacional do município, destaca-se a importância do número de alunos por turma e da carga horária diária. Por fim, obteve-se ineficiência média de 14,7%.

Keywords: Teoria da Firma, Educação, Forma Flexível de Fourier, Eficiência.

Abstract

This paper models a cost function for the Brazilian public elementary school system. Methodologically, we use a Flexible Fourier Form to overcome the biases usually presented in the use of local flexible forms. The results indicate that, *ceteris paribus*, even spending 10% of the GDP with education would not be sufficient to meet the minimum proficiency for 100% of students. We observe that municipalities with higher income, less urban, better income distribution and higher illiteracy rates have higher spending requirements and higher levels of inefficiency. Among the variables that can be managed by the municipality administration are the number of students per class and the school hours per day. Finally, we obtained an average inefficiency rate of about 14.7%.

Keywords: Teoria da Firma, Educação, Forma Flexível de Fourier, Eficiência.

JEL classification: D22, H52, I20, C14

DOI: http://dx.doi.org/10.11606/1413-8050/ea154258

* Secretaria do Tesouro Nacional, MF - carlos.m.castro@tesouro.gov.br
† Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA - geraldo.souza@embrapa.br
‡ Departamento de Economia, Universidade de Brasília - maria.tannuri@gmail.com

Recebido em 10 de setembro de 2015 . Aceito em 12 de julho de 2016.

1 Introdução

O crescimento significativo da oferta de bens públicos em educação ocorrido no passado recente e aquele que se projeta para o Brasil justificam a discussão acerca da estrutura econômica subjacente a tal oferta. Questões alocativas e de eficiência são fundamentais para que as políticas impliquem nos retornos esperados.

O gasto por aluno da educação básica no Brasil mais que dobrou em termos reais de 2005 a 2011, e as metas constantes no novo Plano Nacional de Educação vsilubram passar do nível de 6,1% do PIB em 2011 para 10% no próximo decênio.

Quando comparamos em valores absolutos, o gasto público por aluno no Brasil, apesar de ter avançado significativamente nos últimos anos, ainda está bastante aquém da média dos países da OCDE. Em 2010, a média da OCDE era de US$ 8.382, enquanto o Brasil gastava US$ 2.964. De fato, o alcance dos patamares de gastos por aluno da ordem de US$ 8.000 depende naturalmente também do crescimento da renda do país.

Nesse contexto, algumas questões naturais surgem, na medida em que ampliarmos o volume de recursos aplicados. Qual é a relação existente no caso brasileiro entre o gasto e a performance dos estudantes? Quais são as variáveis que têm mais impacto na estrutura de gastos do Brasil e quais são seus efeitos sobre o resultado do processo educativo? Onde há mais ineficiência na gestão dos recursos?

Nessa linha, pretende-se colaborar com a análise de como funciona a estrutura do gasto em educação no Brasil, através da tentativa de estimar uma função custo para o setor. Com isso, podemos prospectar, por exemplo, se gastos da ordem de 10% do PIB seriam suficientes para alcançar níveis adequados de proficiência.

Em termos metodológicos, utilizamos a técnica de fronteira estocástica considerando uma forma flexível de Fourier para a função custo, onde as firmas seriam os municípios brasileiros. Incluímos preços para trabalho, capital e custeio e consideramos o percentual de alunos proficientes como produto da firma. Adicionalmente, acrescentamos covariáveis relacionadas ao ambiente familiar do aluno e ao ambiente escolar e municipal que tipicamente são consideradas nesses estudos. Dessa forma, temos como estimar quanto custa elevar a performance dos estudantes brasileiros, como comparar a infuência de diversas variáveis sobre o nível de gastos e discutir questões de eficiência.

Os resultados indicam que, na estrutura atual, o alcance da proficiência dos alunos implicaria em um gasto por aluno 277% superior. Isso indicaria que, tudo mais constante, a elevação de gasto de 6% para 10% do PIB não seria suficiente para alcançar a proficiência de todos os alunos brasileiros. Ainda analisando a relação entre custo e produto, observamos que estamos em uma região de deseconomias de escala. Ao comparar variáveis, concluímos a importância das variáveis carga-horária diária e alunos por turma, como instrumentos importantes para racionalização na utilização de recursos. Constatamos que os municípios mais eficientes apresentam piores indicadores de renda e gasto por aluno inferior, de onde inferimos a necessidade de focalização na alocação de recursos para que não haja desperdício e de aprimoramento na gestão para elevação dos níveis de eficiência.

Na sequência, apresentamos a literatura relacionada e o modelo que tenta adaptar os pressupostos da teoria da firma ao contexto de provisão de bens públicos educacionais. A estratégia empírica é abordada amplamente na seção seguinte que inclui a discussão acerca das formas flexíveis locais e globais. Aqui será apresentada a forma flexível de Fourier e a especificação final do estudo. A Seção 5 apresenta a estatística descritiva dos dados e as *proxies* utilizadas para as variáveis consideradas na função custo a ser estimada. Com isso, é possível apresentar a amplitude da base de dados com a qual se trabalhou, detalhando aqueles municípios que ficaram de fora do escopo deste estudo. Por fim, apresentam-se os resultados obtidos e são discutidas as principais perguntas da pesquisa.

2 Literatura

Este trabalho insere-se em uma linha de pesquisa da Economia da Educação que, segundo a classificação de Blaug (1992), preocupa-se em investigar aspectos econômicos dos sistemas educativos. Segundo Waltenberg (2006), o instrumental de análise de um sistema educativo seria mais microeconômico e microeconométrico. Entre os temas tratados, nessa abordagem, está a análise de eficiência na alocação de recursos via funções de produção e funções custos. Ou seja, busca-se mobilizar o ferramental da teoria da firma para que seja possível encontrar a fronteira de possibilidades de produção e assim desenhar políticas. Nessa linha, observa-se um grande número de estudos realizados a partir da publicação do trabalho pioneiro de Coleman et al. (1966) entre os quais se destaca a metanálise realizada por Hanushek (1997).

Não encontramos estudos para o contexto brasileiro que estimasse efetivamente uma função custo nos moldes próximos ao da teoria da firma. Avaliamos que a dificuldade com *proxies* para os preços seria uma das razões para tal escassez. Naturalmente, a dualidade existente entre funções de produção e funções custo nos permitem gerar inferências equivalentes. No entanto, destacamos que, com o objetivo também de discutir eficiência, a função custo nos permite tratar de uma eficiência econômica que engloba aspectos técnicos e alocativos. Em geral, nos modelos de eficiência estocástica, a função custo nos permite avaliar eficiência econômica na presença de *outputs* múltiplos. Intuitivamente nos parece mais natural considerar nossa variável resposta como uma *proxy* de um construto agregando *outputs* múltiplos do que determinante do modelo de produção. Por fim, a ausência de tais estudos nos motiva a esta tentativa.

Não obstante, dentro dessa linha de pesquisa, Machado et al. (2008) analisaram os determinantes do desempenho de alunos em escolas públicas estaduais mineiras. Menezes Filho et al. (2009) estudaram a relação entre gastos e o desempenho educacional de municípios brasileiros. Com o objetivo de medir a qualidade das escolas brasileiras, Curi & Souza (2012) modelaram a relação entre o desempenho dos alunos e características das escolas, dos alunos e outras. Considerando municípios goianos, Rosano-Peña et al. (2012) analisaram a eficiência dos gastos públicos em educação por meio de modelos de análise envoltória dos dados (DEA). Machado & Gonzaga (2007) analisaram o impacto dos fatores familiares sobre a defasagem idade-série para as crianças brasileiras. Com um painel de dados de 1998 a 2005, Soares & Sátyro (2008) concluíram que os insumos escolares têm impacto significativo sobre a taxa de distorção série-idade. de Carvalho & de Sousa (2014) aplicaram um modelo de DEA em três estágios e conclui que, mesmo desconsiderando os fatores ambientais e aleatórios, ainda persistem problemas de gestão nas escolas urbanas das regiões Nordeste e Sudeste brasileiras. Rocha et al. (2013) concluíram, por meio de um modelo de DEA, que os recursos municipais seriam suficientes para alcançar metas do IDEB em 2021 dado o alto nível de desperdício apresentado pelo modelo.

Com isso, buscamos estudos que tentaram modelar uma função custo para sistemas educativos fora do contexto brasileiro. Com o objetivo de estudar economias de escala, Duncombe et al. (1995) estimaram uma função custo para os distritos escolares de Nova Iorque, e Tao & Yuan (2005), para as escolas públicas de Taipei em Taiwan. Considerando um painel de escolas do ensino médio de Nova Iorque, Stiefel et al. (2009) compararam os custos de diversos tipos de organizações escolares. Ajustando uma fronteira estocástica para função custo na forma translog, Gronberg et al. (2012) mostraram que as escolas *charters* geram os mesmos resultados a custos inferiores. Duncombe & Yinger (2005) estimaram o custo adicional que os distritos escolares têm com alunos em desvantagem (baixa renda, baixo rendimento, etc.). Uma análise de metarregressão com base em estudos que estimam função custo é feita por Colegrave & Giles (2008). Como estudos correlacionados destacamos: Burnell (1991) que estudou o impacto da fragmentação dos distritos escolares sobre o gastos, Afonso & St Aubyn (2006) que analisaram a eficiência na provisão de educação secundária em 25 países e Thompson (2011) que estimou funções custo para avaliar o efeito da privatização do serviço de transporte escolar.

Observa-se que aspectos metodológicos diversos surgem na discussão da tecnologia da produção educacional. Uma delas é a destacada por Figlio (1999) acerca das formas funcionais das funções de produção ou de custo utilizadas nesses estudos. Sobre estes, destacam-se os aspectos relacionados às formas funcionais do tipo Translog. Segundo Gallant (1981), há dois métodos frequentemente utilizados em aplicações que buscam aproximar funções de produção: aproximações por séries de Taylor e por séries de Fourier. O trabalho com formas funcionais flexíveis tem utilizado principalmente expansões em Taylor como mecanismo de aproximação. Destaque-se que o teorema de Taylor só se aplica localmente. A aplicabilidade local da aproximação basta para traduzir proposições da teoria da demanda em

restrições sobre os parâmetros do sistema orçamentário aproximado. Contudo, o teorema de Taylor falha na compreensão do comportamento estatístico das estimativas dos parâmetros, bem como nos testes relacionados. White (1980) apresentou a fundamentação dessas críticas. Em função desse problema é que métodos de regressão estatística, segundo Gallant (1981) e Elbadawi et al. (1983), devem essencialmente expandir a função verdadeira em uma série de Fourier geral e não em uma série de Taylor. A forma de Fourier possui uma propriedade de flexibilidade global: ela pode aproximar assintoticamente qualquer função contínua, no sentido da norma de Sobolev. Ela permite inclusive testar a forma translog, que faz parte da sua especificação. Gallant (1982) discutiu os detalhes da utilização da forma de Fourier na estimação de funções custo. E principalmente com base nesse estudo, Chalfant (1984) comparou a utilização da forma de Fourier à forma Box-Cox generalizada, dentro do contexto da agricultura americana.

Com isso, utilizamos neste trabalho a forma flexível de Fourier para estimar a função custo do ensino fundamental público ofertado pelos municípios brasileiros que serão as nossas unidades produtivas. Com base em Kumbhakar & Lovell (2003) e Coelli et al. (2005), estimamos as fronteiras estocásticas assim como em Huang & Wang (2004) para realizarmos análises de eficiência.

3 O Modelo Base

O modelo de produção educacional utilizado neste trabalho está baseado principalmente em Duncombe et al. (1995). Nessa linha, propõe-se algumas adaptações à teoria geral da produção, considerando que trata-se de provisão de serviços públicos, no caso, educação. Em princípio, o processo de produção em uma escola, distrito escolar ou município poderia ser modelado com uma função de produção do tipo:

$$G = f(L, K, T), \tag{1}$$

em que L representaria a força de trabalho (professores e outros), K, o capital, e T, os outros insumos de custeio da firma.

Entre as referidas adaptações, destaca-se a importante discussão acerca do produto e do ambiente no caso de serviços públicos. Como medir produtos que representem quantidade de atividades de qualidade equivalentes? A entrega de uma hora-aula em um dado contexto pode ter qualidade completamente distinta se ocorrese em outra situação. Surge assim a distinção entre serviços diretamente produzidos e aquilo que é de interesse primário do cidadão consumidor. Conforme destacado por Bradford et al. (1969), quando o cidadão vota em um orçamento de segurança pública, por exemplo, ele não está interessado em avaliar se há 10 ou 15 viaturas de polícia, a sua utilidade é afetada principalmente pelo grau de segurança na sua comunidade. De fato, simplificadamente, teríamos:

$$U = U(S, Z) \tag{2}$$

em que Z representa o nível de provisão de outros bens públicos e dos bens privados e

$$S = h(G, E) \tag{3}$$

é obtido pela função h que indica o grau de satisfação do cidadão consumidor com os serviços diretamente produzidos, G, no ambiente E.

Como se observa em grande número de estudos no âmbito educacional, a abordagem em linha com Bradford et al. (1969) é de se medir o produto da escola em termos do desempenho dos alunos em testes padronizados. Este seria um dos interesses primários de quem usufrui do serviço.

Como discutido anteriormente, essa linha de pesquisa tem em Hanushek (1979) e seus trabalhos posteriores uma importante referência em termos da discussão acerca da função de produção para o caso educacional. Segundo Duncombe et al. (1995), entre os fatores ambientais normalmente apontados na literatura (Hanushek (1997)), destacam- se os fatores físicos (P), o background familiar (F) e as características dos estudantes (ST). Ou seja,

$$E = g(P, F, ST). \tag{4}$$

Substituindo 1 e 4 em 3 obtemos a função de produção adaptada, que é a base para gerar a nossa função custo:

$$S = h(G, g(P, F, ST)) \tag{5}$$

Na teoria da produção, a função de produção padrão 1 pode ser resolvida de modo a gerar a função de custo implícita, onde W são os preços:

$$TC = c(G, W) \tag{6}$$

Especificamente, ao resolvermos a Equação 5 para G e substituirmos em 6, obtemos a função custo modificada:

$$TC = c(h^{-1}(S, g(P, F, ST)), W) \tag{7}$$

A Equação 7 nos dá uma forma adaptada onde agregam-se resultados educacionais, variáveis ambientais e preços em uma análise de produção pública.

3.1 As Variáveis

As variáveis utilizadas neste estudo e que comporão a função custo modelada conforme 7 estão descritas em detalhes na seção 5. De fato, além das variáveis de preços (trabalho, capital e custeio), buscamos variáveis ambientais para o caso brasileiro que retratassem o que é observado na literatura. Relativamente aos fatores físicos, incluímos a variável que informa a quantidade de alunos por turma. Quanto ao background familiar, consideramos a infraestrutura da casa do aluno, a participação dos pais na vida escolar e a escolaridade dos pais. Acrescentamos ainda as variáveis carga-horária diária dos alunos e escolaridade dos docentes, pois estão diretamente relacionadas à gestão educacional do município e têm impacto significativo sobre o custo. Por fim, acrescentamos variáveis que tentam retratar o ambiente socioeconômico do município: coeficiente de Gini, taxa de urbanização, taxa de analfabetismo e renda domiciliar per capita. Não encontramos variáveis *proxies* para a variável ST, que reflete a existência de estudantes em condições especiais (aprendizagem, necessidades especiais ou outras).

Destacamos que a unidade produtiva escolhida neste estudo é o município. A disponibilidade de dados foi um dos fatores fundamentais nessa escolha.

4 A Estratégia Empírica

4.1 Forma Flexível de Fourier

Conforme descrito por Gallant (1982), para determinar se uma indústria exibe retornos constantes de escala, se a função de produção é homotética, ou se os insumos são separáveis, uma abordagem comum é especificar uma função custo, estimar os parâmetros usando dados tais como preços e quantidades de insumos e então testar as restrições paramétricas correspondentes a retornos constantes, a tecnologia homotética ou a separabilidade. As chamadas formas funcionais em linha com a Diewert flexibilidade (flexibilidade local) apresentam problemas nessa abordagem, entre os quais se destacam os potenciais vieses na estimação das derivadas e elasticidades.

Segundo White (1980), os coeficientes das formas flexíveis locais seriam corretamente interpretados como estimativas consistentes e não viesadas das derivadas da função verdadeira somente em casos especiais. Ou a função utilizada para aproximar seria da forma da função verdadeira, ou os termos de ordem superior não seriam correlacionados com os termos lineares e quadráticos incluídos. Ele conclui afirmando que a hipótese de regressores ortogonais é "pouco confortável para os econometristas práticos". Como os polinômios de ordem superior em x são variáveis omitidas na interpretação da aproximação por Taylor, as estimativas resultantes não podem ser consideradas ótimas em qualquer sentido aceito como aproximação local. A menos que a forma funcional possa ser considerada como verdadeira, qualquer teste estatístico não pode ser considerado como teste da hipótese de interesse. A rejeição pode ser decorrente da rejeição da forma funcional em vez da proposição econômica que está sendo testada.

É nesse contexto que Gallant (1981) apresenta a forma flexível de Fourier (FFF). Com ela teríamos como construir testes estatísticos não viesados. Estaríamos certos de que é a proposição econômica que está sendo testada e não algum erro de especificação devido à forma funcional utilizada na aproximação da função custo verdadeira. Além disso, estaríamos em condições de estimar de forma não viesada a suas derivadas e, portanto, suas elasticidades. O desenvolvimento teórico dessas ideias, extraídas de Gallant (1982), está exposto no apêndice.

A FFF que garante a ausência de viés citada anteriormente tem como forma geral a expressão:

$$g_k(x|\theta) = u_0 + b'x + \frac{1}{2}x'Cx + \sum_{\alpha=1}^{A}\{u_{0\alpha} + 2\sum_{j=1}^{J}\left[u_{j\alpha}\cos(j\lambda k'_\alpha x) - v_{j\alpha}\sin(j\lambda k'_\alpha x)\right]\} \quad (8)$$

em que $u_0, u_{0\alpha}, u_{j\alpha}, v_{j\alpha,b,C}$ são os parâmetros a serem estimados, sendo b um vetor de ordem 4, no caso que desenvolveremos, e C, uma matriz de ordem 4x4. Denominados de multi-índices, k_α são vetores de ordem 4 e λ, A e J são constantes. A construção dessa forma funcional, incluindo a lógica subjacente aos multi-índices, matrizes, vetores e constantes citados, está detalhada no apêndice.

O ponto inicial para a especificação é a definição do número de parâmetros a ser utilizado na FFF. Elbadawi et al. (1983) discutem tal ponto e indicam que os resultados assintóticos são obtidos fazendo o número de parâmetros depender do tamanho amostral. A especificação dos multi-índices e da ordem da FFF dependem das constantes A e j que aparecem na forma. Esse processo é feito de modo empírico e adaptativo e não é diferente da determinação da defasagem apropriada de um processo ARMA ou de modelos de defasagem distribuída, por exemplo (ver Gallant (1982)).

Chalfant & Gallant (1985), Eastwood & Gallant (1991), Mitchell & Onvural (1996), Huang & Wang (2004) têm sugerido que o número de parâmetros a ser estimado, em um contexto de uma função custo na FFF, seja o número de observações elevado a dois-terços.

Considerando o conjunto de dados do presente estudo que soma 4.600 observações, teríamos que ter no máximo 277 parâmetros na forma. Por outro lado, ao observarmos a FFF descrita em 9, concluímos que ela possui $1 + N + 1 + A(1 + 2J)$ parâmetros. Dado que consideramos três tipos de preços, temos $N = 3$ e, considerando os multi-índices elementares gerados pelo procedimento descrito em Gallant (1982) (e computado pela rotina apresentada por Monahan (1981)), temos as possibilidades para $|k|$, J e A listadas na Tabela 1.

Tabela 1: Nº de Parâmetros

$\lvert k\rvert$	A	J	Nº de Parâmetros
3	56	1	173
3	56	2	285
4	305	1	920

Fonte: Elaboração do autor

Portanto, tais possibilidades nos levaram então a decidir pelos seguintes valores $|k| = 3$ e $J = 1$, dado que temos que ter menos de 277 parâmetros na forma funcional.

Adicionalmente, ao considerar as restrições impostas pelo teorema principal apresentado por Gallant (1982), chegamos à forma da nossa FFF:

$$g(\mathbf{l}, v) = f_1(\mathbf{l}, v) + f_2(\mathbf{l}, v) \tag{9}$$

em que

$$\begin{aligned} f_1(\mathbf{l}, v) = b_0 + b_1 l_1 + b_2 l_2 + b_3 l_3 + b_4 v + c_1 l_1^2 + c_2 l_2^2 + c_3 l_3^2 + c_4 v^2 + \\ c_5 l_1 l_2 + c_6 l_1 l_3 + c_7 l_2 l_3 + c_8 l_1 v + c_9 l_2 v + c_{10} l_3 v \end{aligned} \tag{10}$$

com

$$l_i = \ln(p_i), v = \ln(u)$$

na qual p_1: preço do trabalho p_2: preço do capital p_3: preço do custeio u: *output*

e

$$\begin{aligned} f_2(\mathbf{l}, v) = \quad & -v_1 \sin(\lambda v) - v_2 \sin(\lambda(l_1 - l_3)) - v_3 \sin(\lambda(l_2 - l_3)) \\ & -v_4 \sin(\lambda(l_1 - l_2)) - v_5 \sin(\lambda(l_2 - l_3 + v)) - v_6 \sin(\lambda(l_1 - l_2 + v)) \\ & -v_7 \sin(\lambda(l_1 - l_3 + v)) - v_8 \sin(\lambda(l_2 - l_3 - v)) - v_9 \sin(\lambda(l_1 - l_2 - v)) \\ & -v_{10} \sin(\lambda(l_1 - l_3 - v)) + u_1 \cos(\lambda v) + u_2 \cos(\lambda(l_1 - l_3)) \\ & +u_3 \cos(\lambda(l_2 - l_3)) + u_4 \cos(\lambda(l_1 - l_2)) + u_5 \cos(\lambda(l_2 - l_3 + v)) \\ & +u_6 \cos(\lambda(l_1 - l_2 + v)) + u_7 \cos(\lambda(l_1 - l_3 + v)) + u_8 \cos(\lambda(l_2 - l_3 - v)) \\ & +u_9 \cos(\lambda(l_1 - l_2 - v)) + u_{10} \cos(\lambda(l_1 - l_3 - v)) \end{aligned} \tag{11}$$

4.2 Fronteira Estocástica

A forma funcional 9 é estimada finalmente por meio da técnica de fronteira estocástica. Os modelos de fronteira estocástica foram introduzidos por (Aigner et al. 1977, Meeusen & Van den Broeck 1977, Kumbhakar & Lovell 2003) e apresentam uma boa revisão da metodologia.

A Equação 12 corresponde à especificação do modelo de fronteira estocástica estimado:

$$y = g(\mathbf{l}, v, \mathbf{z}, \mu, \nu) = f_1(\mathbf{l}, v) + f_2(\mathbf{l}, v) + h(\mathbf{z}) + \mu + \nu \tag{12}$$

em que f_1 corresponde à parte Translog da forma flexível de Fourier, f_2 contém os termos trigonométricos característicos da Fourier, h é a combinação linear das covariáveis consideradas, μ é o erro idiossincrático e ν é o termo que caracteriza a ineficiência técnica do sistema produtivo. Destaque-se que y é o log do custo.

Tal modelo nos permite estimar uma fronteira custo-eficiente (representada por unidades que atingem um determinado nível de produto com o menor custo de produção) e calcular índices de ineficiência para as unidades que produzem um custo maior que o custo-eficiente. Nesse modelo, o processo produtivo é afetado por certo nível de ineficiência (captado pelo termo ν) e por choque exógenos aleatórios (captados pelo termo μ). As variáves Z não são propriamente insumos nem produtos do processo produtivo, são consideradas covariáveis que influenciam a estrutura da fronteira de produção, ou seja, a relação entre insumos e produtos (Kumbhakar & Lovell (2003)).

Mais especificamente, temos:

$$h(\mathbf{z}) = b_0 + b_1 z_1 + b_2 z_2 + ... + b_7 z_7 \tag{13}$$

em que z_1= ln(indicador de escolaridade dos docentes) z_2=ln(indicador de carga horária diária dos alunos) z_3=ln(indicador de número de alunos por turma) z_4=ln(indicador de infraestrutura da casa do aluno) z_5=ln(indicador de participação dos pais na vida escolar) z_6=ln(indicador de escolaridade do pai) z_7=ln(indicador de escolaridade da mãe)

Ao estimar a função 9 por meio de uma fronteira estocástica ajustada para função custo, assumimos que:

$$\mu \sim N(0, \sigma)$$

$$\nu \sim Exp(\lambda)$$

em que

$$var(\nu) = \left(\frac{1}{\lambda}\right)^2 = \exp(a_0 + a_8 z_8 + +a_9 z_9 + ... + a_{11} z_{11}) \tag{14}$$

e

z_8=ln(renda domiciliar per capita) z_9=ln(índice de Gini) z_{10}= ln(taxa de urbanização) z_{11}=ln(taxa de analfabetismo)

4.3 Análise Fatorial e Indicadores

Para finalizar a seção relativa à metodologia, apresentamos a técnica que utilizamos para construir três indicadores: o de capital físico das escolas do município, da infraestrutura média da casa das famílias dos alunos do município e da participação dos pais na vida escolar dos alunos.

Para termos uma *proxy* para o preço de capital, foi necessário construir um indicador de quantidade de capital presente nas escolas de cada município. O Censo Escolar da Educação Básica disponibiliza uma série de informações relacionadas à infraestrutura física de cada escola. Especificamente utilizamos informações relacionadas à existência de sala diretor, sala de professores, laboratório de informática, laboratório de ciências, sala de atendimento especializado, quadra de esporte, cozinha, biblioteca, sala de leitura, parque, berçário, sanitário fora da escola, sanitário dentro da escola, sanitário PNE, tv, vídeo, dvd, parabolica, copiadora, retro, impressora, computadores, internet e banda larga.

De maneira idêntica, geramos um indicador para retratar infraestrutura média da casa das famiílias dos alunos do município e outro para medir o nível de participação dos pais na vida escolar dos alunos. Ambas foram consideradas variáveis ambientais importantes nesse contexto. Para a primeira, agregamos informações relativas à presença de tv, radio, dvd, geladeira, freezer, máquina de lavar, computador e de outras facilidades ou bens da família tais como carro, quarto, banheiro dentro da casa. Para a segunda, consideramos informações constantes da Prova Brasil acerca da atuação dos pais no incentivo à execução das tarefas de casa, à leitura, a não faltar às aulas, além da participação em reuniões escolares.

A utilização de técnicas de análise multivariada na formação de indicadores de desenvolvimento, crescimento, pressão do mercado de câmbio, inflação, entre outros, com base na análise de componentes principais e análise fatorial é comum na literatura econômica. Trabalhos recentes nesse contexto são Moreira et al. (2004) e e Souza & Gomes (2015). A definição de indicadores com base em escores obtidos pela análise fatorial tem a desvantagem de basear-se em um sistema em que os pesos das variáveis componentes não são únicos e dependentes de rotações ortogonais dos fatores tipicamente dependentes de percepções subjetivas do investigador. Na análise de componentes principais, a possibilidade de obtenção de pesos negativos torna problemática a interpretação dos escores nas componentes. Nossa abordagem foi escolher como sistema de pesos um conjunto de comunalidades relativas associadas à dispersão de cada variável envolvida no indicador. As comunalidades relativas são intrínsecas à estrutura fatorial e independem de rotação de fatores. Veja Souza e Gomes (2015) para mais detalhes. O coeficiente de correlação múltipla que utilizamos pode ser utilizado como *proxy* para comunalidades como desenvolvido em Moreira, Pinto e Souza (2004). A consideração da transformação de ranks tem três aspectos de fundamental importância. Os ranks captam as tendências de crescimento de cada variável, a medida é robusta relativamente à presença de observações atípicas e heterocedasticidades, a consideração de ranks em análise multivariada permite a utilização de métodos associadas à distribuição normal multivariada em situações mais gerais, finalmente a utilização de ranks permite a agregação de indicadores.

5 Dados

Descrevemos, nesta seção, as variáveis que estarão presentes na função custo a ser estimada. A fonte dos dados e algumas estatísticas básicas são apresentadas. Definimos o universo do estudo e analisamos, por fim, a parcela de municípios ausentes da base final. Importante ressaltar que este estudo tem como escopo o ensino fundamental ofertado pelas escolas públicas municípais e o município é a nossa "unidade produtiva".

O Censo Escolar de 2011 informa que há 16.557.341 matrículas de ensino fundamental na rede municipal. Isso considerando a base mais desagregada disponibilizada pelo INEP/MEC, que apresenta os dados no nível da matrícula[1]. Obtivemos 5.548 municípios com matrículas de ensino fundamental na rede municipal. Isso significa que há no Brasil 15 municípios cujos alunos do ensino fundamental são atendidos apenas pela rede estadual. A menor

[1] Utizamos os códigos 4 a 11, 14 a 21 e 41 da variável FK COD ETAPA ENSINO para caracterizar o aluno do ensino fundamental

rede é a do município paulista de Santana da Ponte Pensa (10 matrículas) e a maior é a do município do Rio de Janeiro (539.214 matrículas).

O Ministério da Educação (MEC) através do Fundo Nacional de Desenvolvimento da Educação (FNDE) coleta informações referentes aos orçamentos de educação dos entes federados. Esses dados estão no Sistema de Informações sobre Orçamentos Públicos em Educação (SIOPE). Especificamente, utilizou-se a despesa paga por município na subfunção "Ensino Fundamental"como *proxy* para o custo. Obteve-se também a segregação da despesa por grupos de despesa, a saber: Pessoal, Outras Despesas Correntes (Custeio) e Investimento.

Construímos uma *proxy* para o preço do trabalho combinando dados da despesa com pessoal, obtidos no SIOPE, com os dados de funcionários presentes no Censo Escolar, que é realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP/MEC). A razão entre a despesa municipal com pessoal na subfunção 'Ensino Fundamental' e a quantidade de funcionários das escolas municipais é o que consideramos como salário anual pago aos profissionais do ensino fundamental das redes municipais no Brasil. A menor remuneração foi registrada no município de Bujaru no Pará e a maior, no município de Porto Alegre.

Observando que o custeio representa uma parcela signficativa da estrutura de custo em tela, foi imprescindível a inserção de uma *proxy* para seu preço. De maneira análoga ao preço de pessoal, consideramos as informações de gasto obtidas no SIOPE[2]. Como proxy para a variável de quantidade, utilizamos o número de salas de aula presentes nas escolas municipais, dado este retirado do Censo Escolar.

De maneira similar aos preços anteriores, construiu-se um preço para o capital através da razão entre o gasto com investimento (SIOPE) e um indicador de quantidade de capital. Tal indicador foi construído a partir de uma série de informações acerca da infraestrutura escolar conforme descrito na respectiva seção de metodologia.

A Prova Brasil é a fonte de dados para construção da variável que consideramos como *output* do processo produtivo na nossa função custo. A Prova Brasil é "uma avaliação censitária envolvendo os alunos da 4ª série/5ºano e 8ªsérie/9ºano do Ensino Fundamental das escolas públicas das redes municipais, estaduais e federal, com o objetivo de avaliar a qualidade do ensino ministrado nas escolas públicas. Participam dessa avaliação as escolas que possuem, no mínimo, 20 alunos matriculados nas séries/anos avaliados, sendo os resultados disponibilizados por escola e por ente federativo"(INEP/MEC). Consideramos especificamente o desempenho na prova de Matemática dos alunos do 5º e 9º anos nas redes municipais. A rigor, a nossa variável final tenta expressar, por município, o percentual de alunos que atingiu um dado patamar de desempenho [3]. Em busca da referência do que seria esse nível ideal de desempenho, adotamos os números apresentados pelo movimento "Todos pela Educação". Tal referência vem sendo utilizada pela sociedade em geral a cada divulgação dos dados da Prova Brasil.

Além das varáveis típicas de uma função custo (custo, preços e *output*), consideramos variáveis ambientais que impactam o processo produtivo por

[2]Considerou-se a diferença entre o gasto total e a soma do gasto de pessoal com o gasto em investimento

[3]Golebiewski (2011) evidencia a utilização dessa proxy como *output* em estudos similares

afetar o ambiente em que são realizadas. São elas: carga horária diária média (Censo Escolar/INEP), percentual de professores com curso superior (Censo Escolar/INEP), Taxa de Analfabetismo (Censo 2010/IBGE), Renda Domiciliar Per Capita (Censo 2010/IBGE), Índice de Gini (Censo 2010/IBGE), Taxa de Urbanização (Censo 2010/IBGE) e indicadores que retratam o a infraestrutura física da residência do aluno e a participação dos pais na vida escolar dele a partir dos dados da Prova Brasil.

Com isso, agregamos as informações por município para todas essas variáveis e chegamos a informações consistentes para 4.600 municípios. Portanto, o nosso estudo terá como universo essa parcela da rede pública municipal brasileira que atua no Ensino Fundamental. Há, portanto, 965 municípios ausentes no nosso estudo, representando, portanto, 17,3% dos 5.565 existentes no país. Eles representam 5,95% da população brasileira e 3,12% da soma dos PIB Municipais do Brasil. Relativamente às matrículas, estamos deixando de fora do estudo cerca de 8,24% das matrículas de ensino fundamental da rede municipal do Brasil.

6 Resultados

Apresentamos, nesta seção, os principais resultados da função custo estimada e as medidas de eficiência decorrentes do modelo de fronteira estocástica ajustado.

6.1 Elasticidades

Com o objetivo de avaliar o efeito marginal de cada um dos preços dos insumos e do produto no custo do processo produtivo, estimamos as elasticidades-preço e a elasticidade-produto da função custo, avaliadas na média de cada uma de suas variáveis (Tabela 3). Todas mostraram-se significativas a 1%.

De maneira similar, obtivemos os efeitos marginais de cada uma das variáveis ambientais, que também mostraram-se muito signficativas, com exceção da taxa de urbanização (Tabela 4).

Relativamente aos preços, os sinais estão de acordo com as expectativas de uma função custo.

A elasticidade-produto é coerente com o que se observa na realidade no que tange ao expressivo custo de se alterar o rendimento dos alunos na atual estrutura produtiva dos municípios brasileiros. Nessa realidade, para que os municípios alcançassem 100% dos alunos com o nível de proficiência proposto, por exemplo, teríamos que ter um custo-aluno 277% superior [4]. Como a elasticidade é superior à unidade, concluímos que, na média, estamos trabalhando em uma região da curva de custo que apresenta deseconomia de escala.

Ao avaliar o ambiente em que o município está inserido, observa-se que os municípios com menos desigualdade de renda e mais ricos implicariam em um custo-aluno maior. É fácil extrair a intuição de que os municípios mais ricos tendem a investir mais em educação.

[4]Elasticidade-produto multiplicada pela variação percentual no produto ((100% − 28,1%)/28,1%). Lembremos que o produto médio (média do percentual de alunos proficientes) é 28,1%).

Tabela 2: Estatísticas Descritivas

Variável	Mínimo	1º quartil	Média	Mediana	3º quartil	Máximo	Desvio-padrão
Matrículas	10	465	2.984	1.128	2.867	539.214	11.472
Despesa	1.625,12	3.542,00	5.359,05	4.782,71	6.445,88	32.694,57	2.528,56
Preço do Trabalho	2.841,69	12.460,85	15.872.31	15.251,70	18.487.91	61.378,76	5.246,09
Preço do Custeio	0	28.738,24	48.195.34	41.356,07	58.891,38	831.239,30	32.387,43
Preço do Capital	0,0006	1,35	8,71	4,24	9,67	700,61	19,47
Output - % alunos proficientes	0	10,5	28,1	28,1	24,4	41,7	100
Matrículas por turma	4,5	17,8	20,5	20,6	23,2	36,5	4,1
Carga Horária Diária	3,6	4,0	4,3	4,2	4,4	9,3	0,5
% Profºs Ensino Superior	0,0	67,6	77,7	84,7	94,1	100,0	21,2
Taxa Analfabetismo	0,9	7,6	15,6	12,6	23,7	47,1	9,8
Renda Domiciliar per capita	95,59	109,40	488,32	463,68	644,35	2.008,98	239,29
Índice de Giini	0,28	0,46	0,50	0,50	0,54	0,80	0,06
Taxa de Urbanização	17,9	49,0	65,6	67,0	83,8	100,0	21,6
Infraestrutura da Casa	454,37	1.770,98	2.626,17	2.665,74	3.487,44	4.749,79	979,82
Participação dos Pais	39,33	1.439,66	2.621,67	2.683,07	3.820,18	5.141,05	1.393,21
Escolaridade do Pai	3	1.348	2.628,51	2.659	3.922,5	5.149	1.485,94
Escolaridade da Mãe	3	1.350	2.622,14	2.641	3.914	5.149	1.482,45

Fonte: Elaboração do autor

Tabela 3: Elasticidades - Preços e Produto

Elasticidade	Coeficiente	Erro-padrão
trabalho	0,167	0,010
capital	0,012	0,003
custeio	0,356	0,008
produto	1,085	0,285

Fonte: Elaboração do autor

Tabela 4: Elasticidades - Ambiente

Elasticidade	Coeficiente	Erro-padrão
carga horária diária	0,369	0,032
alunos por turma	−0,761	0,014
coeficiente de Gini	−0,153	0,021
taxa de urbanização	−0,009	0,007
infraestrutura da casa do aluno	0,246	0,009
participação dos pais	0,060	0,004
escolaridade do pai	−0,017	0,004
escolaridade da mãe	−0,014	0,004
taxa de analfabetismo	0,034	0,008
escolaridade dos docentes	0,043	0,007
renda per capita	0,0002	0,0002

Fonte: Elaboração do autor

Com respeito ao ambiente familiar do aluno, pais mais participativos e com maior infraestrutura em casa implicam em maiores gastos pelo município. Nesse caso, esses resultados podem indicar uma maior capacidade dos pais em exigir mais qualidade na oferta educacional. Na direção contrária, pais com maior escolaridade gerariam menor custo-aluno.

Por fim, as variáveis que estão no âmbito da gestão educacional do município apresentaram sinais coerentes com a expectativa usual. Municípios com maior carga-horária diária, com menos alunos por turma e com maior escolaridade dos seus docentes implicam em maior custo-aluno. Nesse ponto, destacam-se as magnitudes das elasticidades das variáveis "alunos por turma"e "carga horária diária". Por exemplo, podemos afirmar que aumentar a carga-horária média de 4,3 horas para 7 horas (mínimo da educação integral) implicaria em um aumento de 23% no custo. E considerando que a atual média de alunos por turma é de 20,5, obteríamos uma redução da ordem de 35% no custo, se essa média fosse levada para 30 alunos por turma.

6.2 Eficiência

Os municípios têm um custo de ineficiência médio da ordem de 14,7%. Considerando o gasto municipal no ensino fundamental em 2011, isso corresponderia a cerca de R$ 8,9 bilhões. O destaque de eficiência é o município de Morpará na Bahia e o da ineficiência, Xavantina em Santa Catarina. A eliminação dessa ineficiência poderia aumentar a proficiência média em 23%. Tal resultado apresenta-se inferior ao obtido, por exemplo, por Rocha et al. (2013), que foi superior a 40% e por Rosano-Peña et al. (2012), superior a 60% [5].

[5] Além de não utilizarem uma função custo propriamente dita, esses dois estudos utilizaram metodologia distinta: Análise Envoltória de Dados.

A Tabela 5 explicita a relação da ineficiência com as variáveis que a parametrizaram. Nitidamente, observamos que há uma relação direta com maiores níveis de renda, de igualdade de renda e de analfabetismo.

Tabela 5: Coeficientes do Termo de Ineficiência

Variável	Coeficiente	Erro-padrão
renda	0,0030	0,0003
gini	−2,5720	0,3600
tx urb	−0,1510	0,1260
tx analf	0,5790	0,1350
cons	−9,3860	0,6720

Fonte: Elaboração do autor

Apresentamos, a seguir, as características dos 30 municípios mais eficientes (Tabela 6) e dos 30 menos eficientes (Tabela 7). Além disso, comparamos as médias das ineficiências, considerando os 100 municípios menos eficientes com os 100 mais eficientes (Tabela 8).

Observamos que, na média, os mais eficientes apresentam piores indicadores socioeconômicos e menor nível de proficiência. No entanto, o gasto por aluno é bastante inferior nestes, o que parece determinar uma menor ineficiência. Destaque-se ainda que os mais eficientes apresentam menores carga-horária diária, número de alunos por turma e escolaridade dos docentes.

7 Discussão

O modelo construído para estimar uma função custo para o ensino fundamental municipal no Brasil permite comparar o impacto dos preços dos diversos insumos, do produto e das covariáveis no custo-aluno e discutir questões de eficiência. Com isso, alguns direcionamentos de políticas podem ser levantados, considerando os aspectos de custo, resultados e eficiência estudados.

Nesse sentido, as elasticidades-preço mostram que o preço do custeio ordinário tem aproximadamente o dobro do impacto do preço do trabalho e que o preço do capital tem um efeito bastante inferior aos anteriores, considerando as *proxies* construídas para eles. Uma das políticas mais debatidas em educação atualmente refere-se à necessidade de valorização do salário dos professores, razão pela qual, desde 2009, vigora o piso salarial nacional dos professores. Segundo os nossos resultados, o acréscimo de 1% nos salário dos professores e funcionários implicaria em um aumento de 0,16% no custo-aluno.

Quando avaliamos a elasticidade-produto, observamos que, para alcançar 100% dos alunos com a performance mínima, segundo a definição utilizada neste estudo, o custo teria que ser 277% superior. Tal relação nos parece bastante coerente e, combinada com outras ações, permite traçar uma trajetória de crescimento da performance do estudante brasileiro. Considerando que o Plano Nacional de Educação prevê uma trajetória de gasto que sairá de 6% para 10% do PIB, concluímos que tal expansão deve ser combinada com redução das ineficiências e outras ações, pois, senão, não alcançaremos o nível desejado de proficiência. Destacamos que estamos um ponto de deseconomias de escala.

Tabela 6: Os 30 Municípios mais eficientes

Município	Custo	Prof.	Esc.	Carg	Turm	Renda	Gini	Tx.	Infr	Part	Esc.	Esc.	Tx.	Num	Inef.
			Doc.	Hor.				Urb.	Casa	Pais	Mãe	Pai	Analf	Alun	
Unio - PI	2,864	3,6	40,4	4,4	9,3	229,01	0,52	49,2	1,158	1,130	818	805	28,0	6.300	1.022
Nova Cana - BA	3,015	12,1	37,1	3,9	8,0	244,81	0,48	41,0	2,072	332	1611	1,081	31,3	2.777	1.023
Maus - AM	2,347	9,4	43,9	4,2	17,5	238,10	0,65	49,5	1,857	2,323	2992	3,673	9,4	10.417	1.024
Colniza - MT	3,776	15,3	35,5	3,9	11,7	403,85	0,59	56,8	2,657	112	263	642	10,2	40.80	1.025
Barra da Estiva - BA	3,249	21,2	68,6	4,0	11,8	266,34	0,546	49,1	1,887	1,208	859	1,423	15,3	3.622	1.025
Anori - AM	2,468	23,9	51,1	4,0	17,8	226,78	0,60	61,3	1,645	2,600	1712	1,426	17,0	1.798	1,026
Lagoa do Piau - PI	4,179	4,2	57,5	4,2	8,8	242,12	0,47	43,3	1,353	2,106	3995	3,504	26,6	741	1.027
Santana do Acara - CE	1,988	12,9	94,9	4,0	19,2	214,08	0,69	51,3	592	3,514	514	512	27,0	5.702	1.027
So Jos do Cerrito - SC	7,444	23,4	93,7	4,0	4,5	368,29	0,47	26,9	3,008	3,384	1758	2,701	12,7	643	1.028
Beruri - AM	2,384	56,0	38,4	4,0	19,1	189,46	0,66	50,2	1,621	867	591	1,984	27,1	4.161	1.028
Gararu - SE	3,157	4,9	79,7	4,4	9,7	225,64	0,62	24,8	927	1,801	1177	853	29,9	1.564	1.029
Buriti dos Montes - PI	2,866	32,3	84,4	4,3	13,2	211,61	0,57	30,4	1,522	4,818	5006	4,686	29,2	1.397	1.029
Ipu - CE	2,017	12,4	74,6	4,0	21,2	295,10	0,56	63,5	1,374	2,754	742	725	27,6	6.493	1.030
Careiro - AM	3,102	12,0	90,5	4,0	18,4	192,03	0,67	28,8	2,055	181	230	442	13,6	5.122	1.030
Vera Mendes - PI	3,919	1,9	69,1	3,9	8,8	172,40	0,58	32,4	1,698	155	142	90	43,7	820	1.031
Amatur - AM	2,177	2,3	95,5	4,2	18,1	153,52	0,67	52,4	1,513	856	1534	3,436	25,2	1.705	1.031
Paranapu - SP	5,776	58,5	88,9	5,0	20,4	666,85	0,45	89,0	3,110	811	2297	3,603	9,8	265	1.031
Nova Bandeirantes - MT	4,729	19,6	67,5	4,0	9,8	583,71	0,64	34,9	2,237	1,144	1699	2,018	7,3	1.532	1.032
Jaguaquara - BA	2,295	13,7	55,4	4,1	20,7	292,29	0,51	76,2	1,708	2,013	1627	1,644	27,3	7.291	1.032
Itapiranga - AM	2,748	10,1	94,7	3,9	16,0	319,98	0,67	78,6	1,903	3,896	4156	5,044	6,7	1.187	1.032
Monte Alegre de Gois - GO	4,273	36,4	45,8	4,3	9,5	313,81	0,61	40,9	2,415	780	1979	3,649	24,0	762	1.032
Bujaru - PA	1,625	7,0	38,9	4,0	20,7	167,93	0,54	31,5	1,617	208	436	734	14,8	4.811	1.032
So Joo do Triunfo - PR	3,882	35,8	97,8	4,0	15,2	460,84	0,50	29,5	2,391	4,040	1469	1,049	6,6	1.303	1.033
Belm de So Francisco - PE	3,726	14,8	42,8	4,4	9,0	347,18	0,63	62,1	1,725	2,334	4454	2,436	20,0	2.632	1.033
Breves - PA	2,165	11,6	68,1	4,3	25,9	206,39	0,58	50,1	1,835	1,268	2432	3,198	25,8	28.250	1.033
Salitre - CE	3,463	11,6	86,6	3,9	16,8	176,47	0,49	40,5	983	3,310	2187	1,608	39,3	3.847	1.033
Roteiro - AL	3,266	3,1	96,3	4,1	28,5	198,27	0,47	87,6	2,172	108	60	125	37,2	1.624	1.033
Anajs - PA	2,338	4,1	60,4	4,0	19,8	186,20	0,62	38,3	1,474	630	925	1,103	31,2	8.441	1.033
Atalaia do Norte - AM	2,941	21,4	14	4,0	23,9	157,45	0,66	45,5	2,338	4,537	779	1,622	38,6	2.546	1.033
Pauini - AM	2,365	6,2	54,5	4,0	19,2	213,82	0,73	51,0	1,441	2,150	903	3,296	31,1	2.682	1.033

Tabela 7: Os 30 municípios menos eficientes

Município	Custo	Prof.	Esc.	Carg	Turm	Renda	Gini	Tx.	Infr	Part	Esc.	Esc.	Tx.	Num	Inef.
			Doc.	Hor.				Urb.	Casa	Pais	Mãe	Pai	Analf	Alun	
Xavantina - SC	24,591	21,6	70	4,0	11,1	967,36	0,47	27,0	3,374	4,374	1398	4,582	7,5	89	3.228
Ariranha do Iva - PR	18,429	50,0	100	4,0	17,8	420,31	0,41	36,9	2,187	4,779	4666	5,053	15,3	148	2.585
Maca - RJ	15,470	35,9	77,8	4,7	25,4	1.047,01	0,57	98,1	3,505	2,145	4325	4,246	4,1	23.174	2.516
Jate - MS	32,695	5,3	83,3	4,5	17,3	687,26	0,61	46,6	3,047	3,824	3073	2,251	12,6	173	2.442
Campo Bonito - PR	13,801	40,7	100	4,1	18,5	577,25	0,488	58,5	3,381	3,726	3464	3,710	11,6	357	2.412
Itamb - PR	9,635	50,6	100	4,0	19,7	581,21	0,36	94,9	3,206	4,943	5008	4,265	12,8	359	2.390
Esperana Nova - PR	12,191	17,5	100	4,0	20,3	576,61	0,39	38,2	3,341	4,612	5112	4,989	12,1	122	2.263
So Domingos - SC	14,073	60,0	83,3	4,0	20,6	953,33	0,59	66,5	3,971	3,761	4942	3,938	9,2	268	2.256
Itutinga - MG	19,426	50,0	69,2	4,3	16,9	613,41	0,56	70,4	2,523	3,011	91	1,813	6,0	135	2.253
Tocantinpolis - TO	12,526	15,8	85,4	4,3	15,0	410,95	0,53	81,0	1,291	2,802	1525	921	15,7	539	2.247
Muitos Capes - RS	21,576	26,3	100	4,0	20,7	652,56	0,51	32,5	3,679	604	4056	4,635	5,8	124	2.192
Candiota - RS	21,250	0,0	100	4,0	16,3	628,54	0,52	29,6	3,608	357	3045	1,410	6,0	392	2.177
So Joo de Iracema - SP	16,279	60,7	75	9,0	19,6	557,06	0,37	81,6	2,971	4,256	3800	1,465	10,8	98	2.166
Nova Cana Paulista - SP	14,190	60,9	100	5,0	18,0	519,87	0,34	41,6	3,543	4,365	3768	3,778	11,6	90	2.161
Tigrinhos - SC	13,453	45,5	100	4,0	15,0	573,71	0,41	19,5	2,637	4,885	3960	4,736	9,2	120	2.149
Campinas - SP	11,822	27,4	92,1	4,6	27,7	1.320,21	0,58	98,3	3,848	2,076	4207	4,389	3,2	22.470	2.143
Gov. Celso Ramos - SC	8,997	21,7	83	4,0	20,6	777,23	0,44	94,3	3,774	3,926	5.080	5,020	7,0	596	2.126
Ipor - GO	9,144	7,5	100	4,4	20,4	746,24	0,52	91,3	2,817	2,474	3.139	450	10,9	602	2.071
Lins - SP	14,262	41,2	81,5	5,0	23,9	898,97	0,46	98,8	3,777	2,127	1207	2,825	4,2	1387	2.035
Sebastianpolis do Sul - SP	28,273	66,7	100	5,3	15,9	837,93	0,43	77,4	4,359	4,209	4192	3,982	7,0	143	2.030
Nova Aliana do Iva - PR	15,582	54,2	90	4,2	21,0	527,17	0,34	72,7	4,389	4,707	5126	3,487	10,3	108	2.024
Presidente Epitcio - SP	9,912	33,8	91,5	4,8	23,2	664,83	0,49	93,3	3,009	3,809	2747	2,881	6,5	880	2.010
Dona Emma - SC	18,273	70,8	100	4,0	12,0	915,63	0,46	50,2	3,901	4,294	3704	3,487	4,3	72	2.006
Casa Branca - SP	12,731	33,3	87,5	8,9	20,9	700,14	0,51	81,8	3,223	4,491	5.089	5,113	5,4	523	1.999
Guamar - RN	11,944	8,9	96,7	4,5	23,6	395,06	0,53	35,5	2,551	1,506	2582	3,016	24,1	2.152	1.994
Barra Funda - RS	13,866	0,0	84,6	4,1	16,0	755,42	0,39	64,3	2,336	4,669	4666	5,109	5,5	144	1.961
Paulnia - SP	10,760	38,7	94	4,8	29,2	1.145,61	0,49	99,9	3,962	1,922	4705	4,932	3,2	8156	1.917
gua Doce - SC	7,978	42,9	94,7	4,0	24,8	715,20	0,54	49,3	3,867	978	3257	1,902	6,3	545	1.902
Santa Ceclia do Pavo - PR	8,263	35,7	96	4,2	22,8	544,87	0,48	83,8	3,427	3,478	5110	4,063	16,8	254	1.901
Mar de Espanha - MG	8,870	37,1	100	4,3	16,9	594,94	0,40	91,5	2,903	3,340	3363	3,571	8,6	254	1.895
Gabriel Monteiro - SP	11,568	85,7	100,0	5,0	18,3	710,33	0,37	83,3	3893	3900	4714	5046	7,4	146	1.887

Tabela 8: Ineficiência

	100 mais ineficientes	100 menos ineficientes
pib	846.124,00	321.209,00
população	55.153,50	51.254,50
custo	12.413,79	3.390,84
escolaridade do docente	87,77	64,35
carga horária	4,46	4,15
alunos por turma	19,92	17,85
gini	0,48	0,56
taxa urbanização	0,66	0,51
infra casa do aluno	3.289,68	1.927,46
participação dos pais	3.168,33	1.913,13
escolaridade da mãe	3.398,86	2.097,23
escolaridade do pai	3.323,57	2.060,55
taxa de analfabetismo	9,51	22,14
proficiência	36,20	16,61

Consideremos então os resultados obtidos para as variáveis que estão sob a gestão educacional do município. Eles estão entre os que mais podem subsidiar políticas no sentido de alterar a relação entre custo e performance. Por exemplo, podemos afirmar que aumentar a carga-horária média diária de 4,3 horas para 7 horas implicaria em um aumento de 23% no custo. Tal acréscimo de carga horária nos faria alcançar a chamada educação integral e todas as suas externalidades positivas para a performance dos alunos. Por outro lado, a elevação da média de alunos por turma de 20,5 (atual) para 30 provocaria uma redução significativa de 38% no custo, por exemplo.

As elasticidades das variáveis do ambiente socioeconômico do município apresentaram resultados que indicam talvez maiores exigências de gastos nos municípios com mais renda, com menos desigualdade de renda e com mais analfabetos. Infelizmente, isso não parece estar acompanhado de mais eficiência. Tais características também foram observadas ao avaliar os municípios mais ineficientes. A ineficiência média observada de 14,7% pode não parecer muito grande, mas concluímos que ela está presente principalmente nos municípios mais ricos e com menos desigualdade de renda. Estes apresentam melhores resultados educacionais, porém com custo-aluno muito alto. Observemos, por exemplo, que os municípios da região Nordeste aparecem como menos ineficientes, indicando que parece haver uma relação inversa entre renda e eficiência. Tal informação pode ser importante para focalização na alocação de novos recursos.

Dessa forma, concluímos que, na estrutura atual, não parece ser possível alcançar a proficiência de todos os estudantes do ensino fundamental municipal brasileiro, mesmo com gastos da ordem de 10% do PIB. Entre vários outros aspectos, é fundamental combinar políticas que trabalhem com metas para variáveis adequadas (carga-horária, alunos por turma, por exemplo) com possíveis focalizações (municípios com piores indicadores de renda e igualdade de renda) na alocação dos novos recursos e aprimorando a gestão nos mais ineficientes . Isso poderia conduzir a uma elevação significativa da proficiência média dos estudantes. Ratifica-se, por fim, o senso comum de que maiores níveis de gastos sem o respectivo endereçamento dos problemas de eficiência e de gestão das variáveis adequadas implicaria em desperdício óbvio de recursos públicos.

Referências Bibliográficas

Afonso, A. & St Aubyn, M. (2006), 'Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs', *Economic modelling* **23**(3), 476–491.

Aigner, D., Lovell, C. A. & Schmidt, P. (1977), 'Formulation and estimation of stochastic frontier production function models', *Journal of econometrics* **6**(1), 21–37.

Blaug, M. (1992), *The methodology of economics: Or, how economists explain*, Cambridge University Press.

Bradford, D. F., Malt, R. A. & Oates, W. E. (1969), 'The rising cost of local public services: some evidence and reflections', *National Tax Journal* pp. 185–202.

Burnell, B. S. (1991), 'The effect of school district structure on education spending', *Public Choice* **69**(3), 253–264.

Chalfant, J. A. (1984), 'Comparison of alternative functional forms with application to agricultural input data', *American Journal of Agricultural Economics* **66**(2), 216–220.

Chalfant, J. A. & Gallant, A. R. (1985), 'Estimating substitution elasticities with the fourier cost function: Some monte carlo results', *Journal of Econometrics* **28**(2), 205–222.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J. & Battese, G. E. (2005), *An introduction to efficiency and productivity analysis*, Springer.

Colegrave, A. D. & Giles, M. J. (2008), 'School cost functions: A meta-regression analysis', *Economics of Education Review* **27**(6), 688–696.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. & York, R. (1966), 'Equality of educational opportunity', *Washington, dc* .

Curi, A. Z. & Souza, A. P. (2012), 'Medindo a qualidade das escolas: Evidências para o brasil', *XL Encontro Nacional de Economia (Anpec). Salvador* .

de Carvalho, L. D. B. & de Sousa, M. d. C. S. (2014), 'Eficiência das escolas públicas urbanas das regiões nordeste e sudeste do brasil: Uma abordagem em três estágios', *Estudos Econômicos* **44**(4), 649–684.

Duncombe, W., Miner, J. & Ruggiero, J. (1995), 'Potential cost savings from school district consolidation: A case study of new york', *Economics of Education Review* **14**(3), 265–284.

Duncombe, W. & Yinger, J. (2005), 'How much more does a disadvantaged student cost?', *Economics of Education Review* **24**(5), 513–532.

e Souza, G. d. S. & Gomes, E. G. (2015), 'Management of agricultural research centers in brazil: A dea application using a dynamic gmm approach', *European Journal of Operational Research* **240**(3), 819–824.

Eastwood, B. J. & Gallant, A. R. (1991), 'Adaptive rules for seminonparametric estimators that achieve asymptotic normality', *Econometric Theory* **7**(03), 307–340.

Elbadawi, I., Gallant, A. R. & Souza, G. (1983), 'An elasticity can be estimated consistently without a priori knowledge of functional form', *Econometrica: Journal of the Econometric Society* pp. 1731–1751.

Figlio, D. N. (1999), 'Functional form and the estimated effects of school resources', *Economics of Education Review* **18**(2), 241–252.

Gallant, A. R. (1981), 'On the bias in flexible functional forms and an essentially unbiased form: the fourier flexible form', *Journal of Econometrics* **15**(2), 211–245.

Gallant, A. R. (1982), 'Unbiased determination of production technologies', *Journal of Econometrics* **20**(2), 285–323.

Golebiewski, J. A. (2011), 'The literature on education cost functions: An overview'.

Gronberg, T. J., Jansen, D. W. & Taylor, L. L. (2012), 'The relative efficiency of charter schools: A cost frontier approach', *Economics of Education Review* **31**(2), 302–317.

Hanushek, E. A. (1979), 'Conceptual and empirical issues in the estimation of educational production functions.', *Journal of human Resources* **14**(3).

Hanushek, E. A. (1997), 'Assessing the effects of school resources on student performance: An update', *Educational evaluation and policy analysis* **19**(2), 141–164.

Huang, T.-h. & Wang, M.-h. (2004), 'Comparisons of economic inefficiency between output and input measures of technical inefficiency using the fourier flexible cost function', *Journal of Productivity Analysis* **22**(1-2), 123–142.

Kumbhakar, S. C. & Lovell, C. K. (2003), *Stochastic frontier analysis*, Cambridge University Press.

Machado, A. F., Moro, S., Martins, L. & Rios, J. (2008), 'Qualidade do ensino em matemática: determinantes do desempenho de alunos em escolas públicas estaduais mineiras', *Revista da Anpec* **9**(1).

Machado, D. C. & Gonzaga, G. (2007), 'O impacto dos fatores familiares sobre a defasagem idade-série de crianças no brasil', *Revista Brasileira de Economia* **61**(4), 449–476.

Meeusen, W. & Van den Broeck, J. (1977), 'Efficiency estimation from cobb-douglas production functions with composed error', *International economic review* pp. 435–444.

Menezes Filho, N. A., Amaral, L. F. L. E. et al. (2009), A relação entre gastos educacionais e desempenho escolar, Technical report, Insper Working Paper, Insper Instituto de Ensino e Pesquisa.

Mitchell, K. & Onvural, N. M. (1996), 'Economies of scale and scope at large commercial banks: Evidence from the fourier flexible functional form', *Journal of Money, Credit and Banking* pp. 178–199.

Monahan, J. H. (1981), 'Enumeration of elementary multi-indices for multivariate fourier series', *Institute of Statistics, North Carolina State University, Raleigh, Mimeograph Series, n. 1338* .

Moreira, T. B. S., Pinto, M. B. d. P. & Souza, G. d. S. (2004), 'Uma metodologia alternativa para mensuração de pressão sobre o mercado de câmbio', *Estudos Econômicos (São Paulo)* **34**(1), 73–98.

Rocha, F., Duarte, J. et al. (2013), E possível atingir as metas para educação sem aumentar os gastos? uma análise para os municípios brasileiros., Technical report, Textos para Discussão, Secretaria do Tesouro Nacional/MF.

Rosano-Peña, C., Albuquerque, P. H. M. & Marcio, C. J. (2012), 'A eficiência dos gastos públicos em educação: evidências georreferenciadas nos municípios goianos', *Economia Aplicada* **16**(3), 421–443.

Soares, S. & Sátyro, N. (2008), O impacto de infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental: 1998 a 2005, Technical report, Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA).

Stiefel, L., Schwartz, A. E., Iatarola, P. & Chellman, C. C. (2009), 'Mission matters: The cost of small high schools revisited', *Economics of Education Review* **28**(5), 585–599.

Tao, H.-L. & Yuan, M.-C. (2005), 'Optimal scale of a public elementary school with commuting costs:a case study of taipei county', *Economics of Education Review* **24**(4), 407–416.

Thompson, O. (2011), 'The estimated cost impact of privatizing student transportation in minnesota school districts', *Public Choice* **146**(3-4), 319–339.

Waltenberg, F. D. (2006), 'Teorias econômicas de oferta de educação: evolução', *Educação e Pesquisa* **32**(1), 117–136.

White, H. (1980), 'Using least squares to approximate unknown regression functions', *International Economic Review* pp. 149–170.

Apêndice A Forma Flexível de Fourier

Para os nossos objetivos, será suficiente aproximar a função custo em um retângulo do ortante positivo de R^4. Suponhamos que os limites definidos para a formação do retângulo sejam:

$$0 < p_i^L < p_i^H < \infty, \quad i = 1,2,3$$

e

$$0 < u^L < u^H < \infty$$

Essas escolhas são arbitrárias, no entanto, todos os dados e os valores para os quais pretende-se fazer aproximações devem estar neste retângulo. Para este fim, faz-se uso de parâmetros de localização $a_1, a_2, ..., a_4$ e da transformação:

$$l_i = \ln p_i + \ln a_i \quad i = 1,2,3$$

$$v = \ln u + \ln a_{N+1}$$

tal que

$$l_i^L = \ln p_i^L + \ln a_i \quad i = 1,2,...,N$$

$$v^L = \ln u^L + \ln a_{N+1}$$

$$l_i^H = \ln p_i^H + \ln a_i \quad i = 1,2,...,N$$

$$v^H = \ln u^H + \ln a_{N+1}$$

Se $x = (l', v)'$, então a região de aproximação é

$$\Xi = \{x = (l', v)' : l_i^L < l_i < l_i^H, v^L < v < v^H\}$$

que é um retângulo aberto no ortante positivo de R^4. O fecho $\overline{\Xi}$ é também um subconjunto próprio do ortante positivo.

Agora definiremos uma medida de distância, ou seja, uma métrica que nos permita aproximar a função custo e suas derivadas já que temos interesse principalmente no estudo de elasticidades. Do ponto de vista matemático, a métrica que implica essa propriedade é derivada da norma de Sobolev. A função aproximante, nesse contexto, é definida por uma série de Fourier. Nesse ponto, fazem-se necessárias algumas definições que permitirão a caracterização de uma forma funcional aproximante da função custo com base na expansão de Fourier.

Um multi-índice é um vetor de ordem $N+1$ com componentes inteiros. O comprimento do multi-índice é definido como

$$|k|^* = \sum_{i=1}^{N+1} |k_i|$$

O conceito de multi-índice facilita sobremaneira a notação para derivadas parciais mistas de ordem superior.

Nesse contexto, seja λ um multi-índice com componentes não negativos, a derivada parcial de uma função $g(x)$, de ordem λ é denotada por:

$$D^{\lambda}g(x) = \frac{\partial^{|\lambda|^*}}{\partial x_1^{\lambda_1}\partial x_2^{\lambda_2}...\partial x_{N+1}^{\lambda_{N+1}}}g(x)$$

Assumimos que a função $D^{\lambda}g(x)$ é uma função contínua de x quando essa notação for usada. Seja $\mu(x)$ a medida de Lebesgue com suporte no retângulo de interesse Ξ. Seja $g(x)$ uma função definida em Ξ com derivadas parciais de ordem m contínuas. A norma de Sobolev de uma função $g(x)$ de ordem m do tipo p é definida por:

$$||g||_{m,p,\mu} = \left(\sum_{|\lambda|^* \leq m} \int_{\Xi} |D^{\lambda}g(x)|^p d\mu(x) \right)^{1/p}$$

E para $p = \infty$

$$||g||_{m,\infty,\mu} = \sum_{|\lambda|^* \leq m} \sup_{\Xi} |D^{\lambda}g(x)|$$

Segue da definição que a aproximação nessa norma é equivalente à aproximação simultânea da função e de suas derivadas, resultado que se obtém com $p = \infty$. A construção da Forma Flexível de Fourier baseia-se em uma modificação do Teorema de Edmunds e Moscatelli (1977) para a expansão de Fourier, que apresentamos no final deste apêndice. Com esse objetivo considere uma sequência de multi-índices elementares:

$$\mathcal{K}'_{N+1} = \{k_{\alpha} : \alpha = 1, 2, ...A\},$$

de forma que $k_1, k_2, ..., k_{N+1}$ são os vetores elementares e $|k_{\alpha}|^*$ seja não decrescente em α. Defina J como sendo menor inteiro positivo tal que

$$\mathcal{K}'_{N+1} \subset \{jk_{\alpha} : \alpha = 1, 2, ..., A; j = 0, \pm 1, \pm 2, ..., \pm J\}$$

Considere a função aproximante seguinte definida em Ξ:

$$g_k(x|\theta) = a_0 + b'x + \frac{1}{2}x'Cx + \sum_{\alpha=1}^{A}\sum_{j=-J}^{J} a_{j\alpha}e^{ij\lambda k'_{\alpha}x}$$

em que:

$$a_{j\alpha} = \overline{a}_{-j\alpha}, \quad C = -\sum_{\alpha=1}^{A} a_{0\alpha}\lambda^2 k_{\alpha}k'_{\alpha}$$

e a_0, $a_{0\alpha}$, e b são números reais. A barra denota a conjugação complexa e $i = \sqrt{-1}$. O fator de escala λ é calculado como:

$$\lambda = \frac{2\pi - \epsilon}{\max\{x_i^H : i = 1, 2, ..., 4\}},$$

para algum ϵ com $0 < \epsilon < 2\pi$, x_i^H denota l_i^H para $i \leq N$ e v_i^H para $i = N + 1$. Uma escolha razoável é $2\pi - \epsilon = 6$.

Essa função aproximante é uma forma paramétrica formada pela especificação translog adicionada a uma componente associada a uma série de Fourier multivariada. Tal especificação permite dotar a forma translog de propriedades realmente flexíveis no sentido da aproximação uniforme também de suas elasticidades.

O fator λ faz com que $\lambda\Xi = \{\lambda x : x \in \Xi\}$ seja um retângulo com lado não superior a 2π. Esse é um ponto crucial da construção. A locação de $\aleph$ no quadrante positivo através da escolha de a_i é apenas para conveniência das aplicações. O que é essencial é que $\lambda\aleph$ não tenha nenhum lado superior a 2π.

Embora as representações exponenciais complexas sejam mais convenientes na discussão, nas aplicações, as representações em termos de senos e cossenos são mais fáceis de trabalhar. Escrevendo:

$$
\begin{aligned}
a_{0\alpha} &= u_{0\alpha}, \quad \alpha = 1,2,\ldots,A \\
a_{j\alpha} &= u_{j\alpha} + iv_{j\alpha}, \quad \alpha = 1,2,\ldots,A \quad j = 1,2,\ldots,J \\
a_{-j\alpha} &= u_{j\alpha} - iv_{j\alpha}, \quad \alpha = 1,2,\ldots,A \quad j = 1,2,\ldots,J
\end{aligned}
$$

e usando

$$e^{ij\lambda k'_\alpha x} = \cos(j\lambda k'_\alpha x) + i\sin(j\lambda k'_\alpha x),$$

temos que

$$g_k(x|\theta) = u_0 + b'x + \frac{1}{2}x'Cx + \sum_{\alpha=1}^{A}\{u_{0\alpha} + 2\sum_{j=1}^{J}\left[u_{j\alpha}\cos(j\lambda k'_\alpha x) - v_{j\alpha}\sin(j\lambda k'_\alpha x)\right]\} \tag{15}$$

Se $\theta_{(0)} = b = (c', b_{N+1})'$ e $\theta_{(\alpha)} = (u_{0\alpha}, u_{1\alpha}, v_{1\alpha}, \ldots, u_{J\alpha}, v_{J\alpha})'$, então os parâmetros de $g_k(x|\theta)$ são $\theta = (u_0, \theta'_{(0)}, \theta'_{(1)}, \ldots, \theta'_{(A)})'$.

Essa função aproximante é uma forma paramétrica formada pela especificação translog adicionada a uma componente associada a uma série de Fourier multivariada. Tal especificação permite dotar a forma translog de propriedades realmente flexíveis no sentido da aproximação uniforme também de suas elasticidades, que é garantida pelo teorema que se segue, corolário do teorema de Edmunds e Moscatelli (1977, corolário 1). Sua demonstração pode ser vista em Gallant (1982).

Teorema: *Seja $g(x)$ uma função real continuamente diferenciável até a ordem m em um conjunto aberto que contém o fecho de Ξ e que possua homogeneidade linear. Seja*

$$g_k(l,v|\theta) = a_0 + b'x + \frac{1}{2}x'Cx + \sum_{\alpha=1}^{A}\sum_{j=-J}^{J} a_{j\alpha}e^{ij\lambda k'_\alpha x}$$

em que $x = (l', v)'$ está sujeito às seguintes restrições paramétricas:

$$R'_0 : \begin{cases} 1'c = 1 & \text{em que} \quad b = (c', b_{N+1})' \\ a_{j\alpha} = 0 \quad \text{se} \quad 1'r_\alpha \neq 0 & \text{em que} \quad k_\alpha = (r'_\alpha, k_{N+1,\alpha})'. \end{cases}$$

Então $g_k(l,v|\theta)$ possui homogeneidade linear e é possível escolher um vetor de coeficientes $\overline{\theta}_1, \overline{\theta}_2, \ldots, \overline{\theta}_k, \ldots$ tal que para todo p com $1 \leq p \leq \infty$, todo l com $0 \leq l < m$ e todo $\epsilon > 0$,

$$\|g - g_k(\overline{\theta}_k)\|_{l,p,\mu} = o(k^{-m+l+\epsilon}) \quad quando \quad k \to \infty$$

Observe que $a_{j\alpha} = \overline{a}_{-j\alpha}$ *e que* $C = -\sum_{\alpha=1}^{A} a_{0\alpha}\lambda^2 k_\alpha k'_\alpha$*, e que* a_0*,* $a_{0\alpha}$ *e* b *são valores reais.*